

Weekly Report

梅鸿辉

July 22, 2018

1. 预测项目

看了一些文章，和夏老师讨论了一些初步的想法

- 聚合后预测：将数据按时间分块，对每个时间片计算一系列特征值（注），当作多维向量存起来；存储结果用来通过轨迹查询、或者其他knn方法进行检索、匹配和预测
- 结合data cube，在计算每个时间片的时候同时计算它的一些slice，以此来捕捉局部pattern
- 与上述算法相配合的对比可视化和可视分析工具
- 注：特征值包括统计量（聚合函数）以及一些度量（quality metrics）

这个想法还需要进一步的推敲，例如针对的数据是什么？高维时序这样的说法太过笼统。总体来说结合data cube应当是一个方向，毕竟聚合的计算其实是和data cube的应用密切相关。

2. RSATree

整理了相关工作（各种cubes）使用的数据格式、支持的查询方式、系统实现等

整理了一部分当前代码，包括每个函数的作用、使用方式、创建一个完整结构需要的流程等。

R-tree的C/C++化也同样在进行。

此外，和夏老师的聊天中提到的一些问题让我对修改RSATree有了一些新的想法。夏老师说之前的cube类文章感觉缺乏亮点，我认为问题在于太过偏重实现，缺少更高层次的指导思想，因此缺乏后续的发展空间。夏老师也认为case限制太严重。

由此，我觉得我们目前的文章在这方面也没有表达的很好，以至于reviewer也全部盯着性能和对比测试看。或许我们在做好reviewer要求的工作的同时，还需要再对主题进行整理概括和一定程度的升华才行。同时，测试项目也可以针对我们想要突出的特点和亮点进行设计。

3. CHI投稿项目-感知/尺寸相关

看了一些心理学相关文章的研究方法，同时对比了一下他们的使用场景与可视化文章中的差异，还是有很多可以做的内容。可以将他们比较成熟的测试方法以及一些概念用到可视化的使用场景来。

4. 其他

- 签证事宜
- 求职相关：阿里19级应届生内推今日开始，向高翔、关会华了解了一下他们团队的情况，后来又和宁朗聊了一会儿；dataV看起来是可以考虑的选择之一

5. 本周总结

本周平均每天工作约7.5小时，周末约4小时，共约41.5小时。

本周效率有些低，主要原因还是在于时间分配：同时进行三个工作，其中一个的拖延导致了多米诺骨牌般的效应。这方面还是我必须要加强的一个方面。

Papaer Reading

5.1 perception-density

心理学和视觉的一些文章，看看他们的试验方法

[1] F. Yang, L. Harrison, R. A. Rensink, S. Franconeri, and R. Chang, “Correlation Judgment and Visualization Features: A Comparative Study,” *IEEE Trans. Vis. Comput. Graph.*, vol. X, no. X, 2018.

定义了很多散点图的度量（例如点到回归线的平均距离来表达相关性程度）

[2] H. Lee, J. Baek, and S. C. Chong, “Perceived magnitude of visual displays: Area, numerosity, and mean size,” *J. Vis.*, vol. 16, no. 3, pp. 12, 1–11, 2016.

与显示尺寸相关的一个实验

[3] E. Zimmermann and G. R. Fink, “Numerosity perception after size adaptation,” *Sci. Rep.*, vol. 6, pp. 1–7, 2016.

这个的试验方法很有趣，通过视觉误差（误导）来改变变量

[4] J. Bell, A. Manson, M. Edwards, and A. I. Meso, “Numerosity and density judgments: Biases for area but not for volume,” *J. Vis.*, vol. 15, no. 2, pp. 18–18, 2015.

看了下与立体视觉有关，不太相关，pass

[5] W. LIU, Z.-J. ZHANG, and Y.-J. ZHAO, “Numerosity Adaptation Effect on the Basis of Perceived Numerosity,” *Acta Psychol. Sin.*, vol. 44, no. 10, pp. 1297–1308, 2013.

方法和其他的类似，不过提到了这个方法的学名：Point of Subjective Equality (PSE)

[6] M. E. Doherty and R. B. Anderson, “Variation in scatterplot displays,” *Behav. Res. Methods*, vol. 41, no. 1, pp. 55–60, 2009.

各种论文中出现的散点图和他们的参数选择的收集

5.2 Others

[1] D. Liu, P. Xu, and L. Ren, “TPFlow : Progressive Partition and Multidimensional Pattern Extraction for Large-Scale Spatio-Temporal Data Analysis.”

冬煜他们的VAST best paper。通过交互对数据进行划分，然后通过张量分解寻找时空数据中的潜在模式。思路新颖、实现简单，很多思路和套路值得借鉴

[2] M. Behrisch et al., “Quality Metrics for Information Visualization,” vol. 37, no. 3, 2018.

quality metrics的度量，和我研究的visualization recommendation、散点图等都有关系

此外，我寻找了一些time series classification/prediction的文章，粗略查看其中也有很多利用聚合量（主要是平均值）来提高预测准确度的，准备下周仔细阅读并整理其中用到数据抽象、模型、算法等

计划-已完成

TASK	DESCRIPTION	NOTE
大黑书修订	初稿	可能需要进一步修订
研 究 生 课 程PPT修改	初稿	可能需要进一步修订

计划-短期

TASK	DESCRIPTION	SCHEDULE
大黑书修订	继续修改	
研究生课程PPT修改	继续修改	
尺寸感知项目		写motivation等，设计实验
专利（两篇）	完成(1/2)	
CHI投稿		

计划-中期

TASK	DESCRIPTION	SCHEDULE
RSATree代码重构	后端C++化	进行中
RSATree对比实验	跑通其它相关的开源项目	尽快开始进行
RSATree论文	修订后投TVCG	上述工作完成后开始

计划-长期

TASK	DESCRIPTION	SCHEDULE
毕业论文	目前定位为可视设计方向	开始考虑一下整体构思

Works Progresses

TASK	PROGRESS	TODO	ISSUES	DATE
RSATree	修订投TVCG	整理代码、跑通对比项目		
RSATree专利	(1/2)			下周
大黑书修订				
VisEvo		idea evaluation		
电子学报	已进入最后阶段			
ECharts论文	完成proof	等待最终发布		
分辨率自适应可视化		设计实验		